

The Locality of Distributed Symmetry Breaking

Leonid Barenboim* Michael Elkin† Seth Pettie‡ Johannes Schneider§

April 6, 2012

Abstract

We present new bounds on the locality of several classical symmetry breaking tasks in distributed networks. A sampling of the results include

1. A randomized algorithm for computing a maximal matching (MM) in $O(\log \Delta + (\log \log n)^4)$ rounds, where Δ is the maximum degree. This improves a 25-year old randomized algorithm of Israeli and Itai that takes $O(\log n)$ rounds and is *provably optimal* for all $\log \Delta$ in the range $[(\log \log n)^4, \sqrt{\log n}]$.
2. A randomized maximal independent set (MIS) algorithm requiring $O(\log \Delta \sqrt{\log n})$ rounds, for all Δ , and only $2^{O(\sqrt{\log \log n})}$ rounds when $\Delta = \text{poly}(\log n)$. These improve on the 25-year old $O(\log n)$ -round randomized MIS algorithms of Luby and Alon et al. when $\log \Delta \ll \sqrt{\log n}$.
3. A randomized $(\Delta+1)$ -coloring algorithm requiring $O(\log \Delta + 2^{O(\sqrt{\log \log n})})$ rounds, improving on an algorithm of Schneider and Wattenhofer that takes $O(\log \Delta + \sqrt{\log n})$ rounds. This result implies that an $O(\Delta)$ -coloring can be computed in $2^{O(\sqrt{\log \log n})}$ rounds for all Δ , improving on Kothapalli et al.'s $O(\sqrt{\log n})$ -round algorithm.

We also introduce a new technique for reducing symmetry breaking problems on low arboricity graphs to low degree graphs. Corollaries of this reduction include MM and MIS algorithms for low arboricity graphs (e.g., planar graphs and graphs that exclude any fixed minor) requiring $O(\sqrt{\log n})$ and $O(\log^{3/4} n)$ rounds w.h.p., respectively.

*Dept. of Computer Science, Ben-Gurion University. Email: leonidba@cs.bgu.ac.il.

†Dept. of Computer Science, Ben-Gurion University. Email: elkinm@cs.bgu.ac.il.

‡Dept. of Electrical Engineering and Computer Science, University of Michigan. Email: pettie@umich.edu.

§Computer Engineering and Networks Laboratory, ETH Zurich. Email: schneider@tik.ee.ethz.ch.

1 Introduction

Breaking symmetry is one of the central themes in the theory of distributed computation. At initialization the nodes of a distributed system are assumed to be in the same state (but with distinct node IDs), yet to perform any computation the nodes frequently must take different roles, that is, they must somehow break their initial symmetry. In this paper we study three of the classical symmetry breaking tasks in Linial’s \mathcal{LOCAL} model [20]: computing *maximal matchings* (MM), *maximal independent sets* (MIS), and $(\Delta + 1)$ -coloring, where Δ is the maximum degree.¹ In the \mathcal{LOCAL} model each node of the input graph G hosts a processor, which is aware of its neighbors and an upper bound on the size of the graph. The computation proceeds in synchronized rounds in which each processor sends one unbounded message along each edge, which may be different for each edge. *Time* is measured by the number of rounds; local computation is free. At the end of the computation each node must report whether it is in the MIS, or which incident edge is part of the MM, or its assigned color. See [26, Ch. 1-2] for an extensive discussion of distributed models.

Prior Work. The vertex coloring, MM, and MIS problems have been the subject of intensive research since the mid-1980s [1, 2, 3, 9, 10, 11, 19, 20, 22, 24, 28, 31]. In 1986 Israeli and Itai [11] devised a randomized algorithm that computes a MM in $O(\log n)$ time with high probability,² and the same year Luby [22] and Alon, Babai, and Itai [1] independently proposed $O(\log n)$ -time randomized MIS algorithms, which can also be used to compute $(\Delta + 1)$ -colorings in $O(\log n)$ time. These are the fastest known algorithms for MM and MIS on *general* graphs. It was recently shown that $(\Delta + 1)$ -coloring can be computed faster [31], in $O(\log \Delta + \sqrt{\log n})$ time w.h.p. Kuhn, Moscibroda, and Wattenhofer [18] proved that there exist n -vertex graphs with maximum degree $2^{\Theta(\sqrt{\log n})}$ on which any algorithm for MM or MIS (even randomized) requires $\Omega(\sqrt{\log n})$ time. This implies a lower bound of $\Omega(\min\{\log \Delta, \sqrt{\log n}\})$ for these and many other problems. We will henceforth refer to this result as the *KMW bound*.

For deterministic algorithms the situation looks quite different. The fastest MIS and $(\Delta + 1)$ -coloring algorithms on general graphs run in $O(\Delta + \log^* n)$ time [3, 16] and $2^{O(\sqrt{\log n})}$ time [2, 28], whereas the fastest MM algorithms on general graphs run in $O(\Delta + \log^* n)$ time [27] and $O(\log^4 n)$ time [10]. For certain graph classes the bounds cited above can be improved. Barenboim and Elkin [3] showed that on graphs of arboricity λ , MM and MIS can be computed in time $O(\log n / \log \log n)$ for λ sufficiently small (λ must be less than $\log^{1-\epsilon} n$ for MM and less than $\log^{1/2-\epsilon} n$ for MIS.) We believe *arboricity* is an important graph parameter as it robustly captures the notion of *sparsity* without imposing any strict constraints. A graph has arboricity λ if its edge set can be covered by λ forests, or equivalently [25], if every subgraph has density less than λ .³ Lenzen and Wattenhofer [19] gave a randomized MIS algorithm for unoriented trees ($\lambda = 1$) running in $O(\sqrt{\log n} \cdot \log \log n)$ time.⁴ The MM and MIS problems on graphs of bounded growth have also been studied recently [9], culminating in an algorithm [31] running in $O(\log^* n)$ time on

¹The MM problem is to compute a maximal set of vertex-disjoint edges. The MIS problem is to compute a maximal set of vertices, no two of which are adjacent. The $(\Delta + 1)$ -coloring problem is to assign colors from the palette $\{1, \dots, \Delta + 1\}$ such that no edge is monochromatic.

²With *high probability* (w.h.p.) means with probability $1 - 1/n^c$, for an arbitrarily large fixed constant c . All randomized algorithms cited in the paper finish their computation in the stated time bound, w.h.p.

³Note that many sparse graph classes have $\lambda = O(1)$, such as planar graphs, graphs avoiding a fixed minor, bounded genus graphs, and graphs of bounded degree or tree/clique/pathwidth.

⁴The claimed time was $O(\sqrt{\log n \log \log n})$ but there was a flaw in the analysis. See Section 7.

Maximal Matching			Maximal Independent Set		
Citation	Running Time	Graphs	Citation	Running Time	Graphs
II86 [11]	$\log n$	general	L,ABI [1, 22]	$\log n$	general
HKP01 [10]	$\log^4 n$ (Det.)	general	PS95 [28]	$2^{O(\sqrt{\log n})}$ (Det.)	general
PR01 [27]	$\Delta + \log^* n$ (Det.)	general	BE,K [4, 16]	$\Delta + \log^* n$ (Det.)	general
BE08 [3]	$\frac{\log n}{\log \log n}$ (Det.)	$\lambda < \log^{1-\epsilon} n$	BE08 [3]	$\frac{\log n}{\log \log n}$ (Det.)	$\lambda < \log^{1/2-\epsilon} n$
BE08 [3]	$\lambda + \log n$ (Det.)	general	BE09 [4]	$\lambda \sqrt{\log n}$ (Det.)	$\lambda > \sqrt{\log n}$
			SW08 [31]	$\log^* n$ (Det.)	bounded growth
			LW11 [19]	$\sqrt{\log n} \log \log n$	trees ($\lambda = 1$)
L87 [20]	$\Omega(\log^* n)$ (Rand./Det.)	$\Delta \geq 2$	L87 [20]	$\Omega(\log^* n)$ (Rand./Det.)	$\Delta \geq 2$
KMW04,10, [18]	$\Omega(\sqrt{\log n})$ (Rand./Det.) $\Omega(\log \Delta)$	general	KMW04,10, [18]	$\Omega(\sqrt{\log n})$ (Rand./Det.) $\Omega(\log \Delta)$	general
This paper	$\log \Delta + \log^4 \log n$	general	This paper	$\log \Delta \sqrt{\log n}$	general
	$\min \begin{cases} \log \lambda + \sqrt{\log n} \\ \log \Delta + \lambda + \log \log n \end{cases}$	all λ		$2^{O(\sqrt{\log \log n})}$	$\Delta = \log^{O(1)} n$
	$\log \Delta + \frac{\log \log n}{\log \log \log n}$	$\lambda < \log^{1-\epsilon} \log n$		$\min \begin{cases} \log \lambda \sqrt{\log n} + \log^{3/4} n \\ \log \Delta + \lambda + \log \log n \end{cases}$	all λ
		$\log \Delta (\log \Delta + \frac{\log \log n}{\log \log \log n})$		$\lambda \leq \log^{1/2-\epsilon} \log n$	
		$\min \begin{cases} \sqrt{\log n} \log \log n \\ \log \Delta \log \log \Delta + \frac{\log \log n}{\log \log \log n} \end{cases}$		trees ($\lambda = 1$)	
		$\log \Delta \log \log \Delta + 2^{O(\sqrt{\log \log n})}$		girth > 6	

Vertex Coloring			
Citation	Colors	Running Time	Notes
L87 [20]	$\Delta + 1$	$\Omega(\log^* n - \log^* \Delta)$	all $\Delta \geq 2$
BE09,K09 [4, 16]		$\Delta + \log^* n$ (Det.)	
PS95 [28]		$2^{O(\sqrt{\log n})}$ (Det.)	
SW10 [32]		$\log \Delta + \sqrt{\log n}$	
This paper		$\log \Delta + 2^{O(\sqrt{\log \log n})}$	
		$\log \Delta + \min \begin{cases} \lambda^{1+\epsilon} + \log \lambda \log \log n \\ \lambda + \lambda^\epsilon \log \log n \end{cases}$	all λ , fixed $\epsilon > 0$
This paper	$\Delta + O(\lambda)$	$\log \Delta + \lambda^\epsilon \log \log n$	all λ , fixed $\epsilon > 0$
This paper	$\Delta + \lambda^{1+\epsilon}$	$\log \Delta + \log \lambda \log \log n$	all λ , fixed $\epsilon > 0$
KSOS [15]	$O(\Delta)$	$\sqrt{\log n}$	
SW10 [32]		$k + \log^* n$	$k \leq \log \log n, \Delta > \log^{1+1/k} n$
BE10 [5]		$\Delta^\epsilon \log n$ (Det.)	fixed $\epsilon > 0$
This paper		$2^{O(\sqrt{\log \log n})}$	
BE10 [5]	$O(\lambda)$	$\lambda^\epsilon \log n$ (Det.)	fixed $\epsilon > 0$
BE10 [5]	$\Delta^{1+\epsilon}$	$\log \Delta \log n$ (Det.)	fixed $\epsilon > 0$
BE10 [5]	$\lambda^{1+\epsilon}$	$\log \lambda \log n$ (Det.)	fixed $\epsilon > 0$
SW10 [32]	$\Delta \log^{(k)} n$	k	$k \leq \log^* n, \Delta > \log^{1+1/k} n$
L92 [21]	Δ^2	$\log^* n$ (Det.)	

Figure 1: A summary of upper and lower bounds for MM, MIS, and vertex coloring. Here Δ is the maximum degree and λ the arboricity. All running times are randomized (w.h.p.) unless noted otherwise.

this graph class. The $\log^* n$ term in the running times cited above cannot be improved, by Linial's lower bound [20].

Faster coloring algorithms are known if one allows more than $(\Delta + 1)$ colors. Linial [20] devised a deterministic $O(\Delta^2)$ -coloring algorithm requiring $\log^* n + O(1)$ time, which was improved to $\frac{1}{2} \log^* n + O(1)$ by Szegedy and Vishwanathan [30]. Kothapalli et al. [15] gave a randomized $O(\Delta)$ -coloring algorithm running in $O(\log \Delta + \sqrt{\log n})$ time, for all Δ , and Schneider and Wattenhofer [31] devised a randomized $O(\Delta + \log^{1+1/k} n)$ -coloring algorithm running in time $O(k + \log^* n)$. Barenboim and Elkin [5] showed that $\Delta^{1+\epsilon}$ -coloring can be computed in $O(\log \Delta \cdot \log n)$ time deterministically, for any $\epsilon > 0$. Graphs of bounded arboricity λ were shown [3, 5] to be amenable to faster coloring algorithms. In particular, $\lambda^{1+\epsilon}$ -coloring can be computed deterministically in $O(\log \lambda \cdot \log n)$ time and $O(\lambda)$ -coloring in $O(\lambda^\epsilon \cdot \log n)$ time [5], for any fixed $\epsilon > 0$.

Our Results. We give a new randomized MM algorithm running in $O(\log \Delta + \log^4 \log n)$ time, improving the 25-year old bound of $O(\log n)$ [11] and $O(\Delta + \log^* n)$ [27]. According to the KMW lower bound our algorithm is provably optimal whenever $\log \Delta \in [\log^4 \log n, \sqrt{\log n}]$.

We give a randomized MIS algorithm running in $O(\log \Delta \cdot \sqrt{\log n})$ time, improving the 25-year old $O(\log n)$ -time algorithms of Luby [22] and Alon, Babai, and Itai [1] when $\log \Delta \ll \sqrt{\log n}$. If $\Delta = (\log n)^{O(1)}$ we provide an even faster algorithm running in $2^{O(\sqrt{\log \log n})}$ time. These are the first general MIS algorithms running in sublogarithmic time for such a wide range of Δ .

For vertex coloring we give a $(\Delta + 1)$ -coloring algorithm running in $O(\log \Delta + 2^{O(\sqrt{\log \log n})})$ time, improving an $O(\log \Delta + \sqrt{\log n})$ -time algorithm of Schneider and Wattenhofer [32]. As a result of this, we can now compute $O(\Delta)$ -colorings in $2^{O(\sqrt{\log \log n})}$ time for all Δ , improving the $O(\sqrt{\log n})$ -time algorithm of Kothapalli et al. [15].

As noted above, Δ is a significantly more sensitive graph parameter than the arboricity λ . We give a new technique for reducing the symmetry breaking problems on low arboricity graphs to low degree graphs, which is of independent interest. As direct corollaries, our reduction shows that MM and MIS can be solved in $O(\log \lambda + \sqrt{\log n})$ time and $O(\log \lambda \cdot \sqrt{\log n} + \log^{3/4} n)$ time, resp., on graphs with arboricity λ . In particular, for planar graphs (and, more generally, for graphs that exclude any fixed minor), our MM and MIS algorithms require only $O(\sqrt{\log n})$ and $O(\log^{3/4} n)$ time, respectively.

When λ is very small we give several algorithms that are faster for certain ranges of Δ . For example, when $\lambda = O(1)$ our MM, $(\Delta + 1)$ -coloring, and MIS algorithms run in time, respectively, $O(\log \Delta + \frac{\log \log n}{\log \log \log n})$, $O(\log \Delta + \log \log n)$, and $O(\log^2 \Delta + \log \log n)$. These time bounds are exponentially faster, as a function of n , over previous deterministic algorithms [3, 5]. For the special case of trees ($\lambda = 1$) we give an even faster MIS algorithm whose running time is the minimum of $O(\sqrt{\log n \log \log n})$ and $O(\log \Delta \log \log \Delta + \frac{\log \log n}{\log \log \log n})$, which improves on [19].

See Figure 1 for a comparison of our results with prior work.

Technical Summary. All of our algorithms take the following two-phase approach. In Phase I we use some iterated randomized procedure that, with high probability, finds a large partial solution (a matching, independent set, or partial coloring) that effectively breaks the global problem into a collection of disjoint subproblems with $\text{poly}(\log n)$ size or $O(\sqrt{\log n})$ diameter. In Phase II we solve each subproblem using the best available deterministic algorithm. It is for this reason that our running times are usually *exponentially* faster in terms of n than the best deterministic algorithms, e.g., a $2^{O(\sqrt{\log n})}$ bound becomes $2^{O(\sqrt{\log \log n})}$, a $\frac{\log n}{\log \log n}$ bound becomes $\frac{\log \log n}{\log \log \log n}$ and so on. This

strategy has been used in other contexts, for example, in Beck's [6] algorithmic approach to the Lovász Local Lemma, the local hypergraph coloring algorithms of Rubinfeld et al. [29], and the $O(\Delta)$ -coloring algorithm of Kothapalli et al. [15]. The main technical difficulty is in the analysis of Phase I's iterated randomized procedure.

Our analyses often bound the running time in terms of Δ , which can be significantly larger than the arboricity λ . We give a new reduction that, roughly speaking, reduces the maximum degree to $\lambda \cdot 2^{\log^\epsilon n}$ in $O(\log^{1-\epsilon} n)$ time, for any $\epsilon \in (0, 1)$. This allows us to achieve sublogarithmic (in n) running times using algorithms that depend logarithmically on Δ .

Organization. Section 2 introduces some terminology and notation. Our MM, MIS, and $(\Delta + 1)$ -coloring algorithms are presented in Sections 3–5. Section 6 presents a reduction from graphs with small arboricity to small degree. In Section 7 we give an improved MIS algorithm for trees ($\lambda = 1$) and girth > 6 graphs, and correct a small flaw in the analysis of [32]. Appendix A restates some standard concentration inequalities.

2 Preliminaries

All logarithms are base 2 unless noted otherwise. The input graph is $G = (V, E)$. For any $V' \subseteq V$, let $G(V')$ be the subgraph of G induced by V' . Let $\Gamma_G(v) = \{u \mid (v, u) \in E\}$ and $\deg_G(v) = |\Gamma_G(v)|$ be the neighborhood of v in G and its cardinality. Let $\hat{\Gamma}_G(v) = \Gamma_G(v) \cup \{v\}$ be the neighborhood including v . Let $\Delta = \Delta(G) = \max_{v \in V} \deg_G(v)$ be the maximum degree. Let $\text{dist}_G(u, v)$ be the length of the shortest path (i.e., distance) between u and v in G . The *diameter* of G is $\max_{u, v \in V} \text{dist}_G(u, v)$ and the *weak diameter* of a subgraph $G(V')$ is the maximum distance between V' -vertices with respect to G , that is, $\max\{\text{dist}_G(u, v) \mid u, v \in V'\}$. In a directed graph the indegree (outdegree) of v is the number of edges directed to (from) v , and the degree of v is the sum of its in and outdegree. A *forest* is an acyclic graph. An *oriented forest* is a directed forest in which each non-root has outdegree 1; a *pseudoforest* is a directed graph in which all vertices have outdegree 0 or 1. The *arboricity* $\lambda = \lambda(G)$ is the minimum number of edge-disjoint forests that cover E . By the Nash-Williams theorem [25], $\lambda(G)$ is equivalently expressed as $\max \left\{ \left\lceil \frac{|E(G(V'))|}{|V(G(V'))|-1} \right\rceil : V' \subseteq V, |V'| \geq 2 \right\}$, i.e., roughly the edge-density of any subgraph $G(V')$ with at least 2 vertices. Many other measures of graph sparsity are, for our purposes, equivalent to λ , such as *degeneracy* or *k-core number*.

In our analyses we use several standard concentration inequalities due to Chernoff, Janson, and Azuma-Hoeffding. See Appendix A or [7] for statements of these theorems.

3 An Algorithm for Maximal Matching

The Match procedure given below is a generalized version of the Israeli-Itai MM algorithm [11]. (See also [33].) It is given two vertex sets U_1, U_2 (not necessarily disjoint) and a matching M , and returns a matching on $U_1 \times U_2$ vertex-disjoint from M .

Match(U_1, U_2, M)

1. Initialize *directed* graphs $F_1 = (U_1 \cup U_2, \emptyset)$ and $F_2 = (U_1 \cup U_2, \emptyset)$.
2. Each $v \in U_1 \setminus V(M)$ chooses a neighbor $u \in U_2 \setminus V(M)$ uniformly at random and includes (v, u) in $E(F_1)$. (Note: F_1 is a pseudoforest.)

3. Each $u \in U_2$ with $\text{indeg}_{F_1}(u) > 0$ chooses the $v' \in \{v : (v, u) \in E(F_1)\}$ with maximum node ID and includes (v', u) in $E(F_2)$. (Note: F_2 consists of directed paths and cycles.)
4. If $\text{deg}_{F_2}(v) = 2$ then v chooses a bit $b(v) \in \{0, 1\}$ uniformly at random. Otherwise $b(v) = 0$ (respectively, 1) if v is at the beginning (resp., end) of a path in F_2 .
5. Return the matching $\{(v, u) \in E(F_2) : b(v) = 0 \text{ and } b(u) = 1\}$.

Note that U_1 and U_2 are allowed to contain matched vertices since these are specifically excluded in step 2. Phase I of our maximal matching algorithm consists of a sequences of $\Theta(\log \Delta)$ stages. In the pseudocode below M_i is the matching M just before stage i , $V_i = V \setminus V(M_i)$ is the set of unmatched vertices before stage i , and deg_i and Γ_i are the degree and neighborhood functions w.r.t. $G(V_i)$. Define the parameters δ_i , τ_i , and ν_i as

$$\delta_i = \frac{\Delta \sqrt{c_1 \ln n}}{\rho^i}, \quad \tau_i = \frac{2\Delta}{\rho^i \sqrt{c_1 \ln n}}, \quad \nu_i = \frac{\Delta^2}{\rho^{2i}} = \frac{\delta_i \tau_i}{2}$$

where c_1 is a sufficiently large constant and $\rho \approx 1.06$ a constant to be determined precisely later. Define $V_i^{\text{lo}} = \{v \in V_i : \text{deg}_i(v) \leq \tau_i\}$ and $V_i^{\text{hi}} = \{v \in V_i : \text{deg}_i(v) > \delta_i\}$ to be the low and high degree vertices at the beginning of stage i . In each stage i we supplement the current matching M_i first with a matching on $V_i^{\text{lo}} \times V_i^{\text{hi}}$, then with a matching on V_i .

Phase I: Initialize $M_0 \leftarrow \emptyset$ and execute stages $0, \dots, c_2 \log \Delta - 1$.

Stage i :

1. $M_{i+1} \leftarrow M_i \cup \text{Match}(V_i^{\text{lo}}, V_i^{\text{hi}})$
2. $M_{i+1} \leftarrow M_{i+1} \cup \text{Match}(V_i, V_i)$

Phase II: Let \mathcal{C} be the connected components induced by $V_{c_2 \log \Delta}$ with size at most $\log^9 n$. Deterministically compute a maximal matching $M(C)$ on each $C \in \mathcal{C}$ and return $M_{c_2 \log \Delta} \cup \bigcup_{C \in \mathcal{C}} M(C)$.

The algorithm always returns a matching. If, at the beginning of Phase II, \mathcal{C} contains all connected components on $V_{c_2 \log \Delta}$ then the returned matching is clearly maximal. Thus, our goal is to show that with high probability, after Phase I there is no connected component of unmatched vertices with size at most $\log^9 n$. In the proof below $\text{deg}(S)$ is short for $\sum_{u \in S} \text{deg}(u)$ for $S \subset V$.

Lemma 3.1 *Let l be any index for which $\tau_l > c_3 \ln n$ for a sufficiently large constant c_3 . Then for all $i \in [0, l]$, $\text{deg}_{i+1}(v) \leq \delta_i$ and $\text{deg}_{i+1}^{(2)}(v) \stackrel{\text{def}}{=} \text{deg}_{i+1}(\Gamma_{i+1}(v)) \leq \nu_i$ with probability $1 - 1/\text{poly}(n)$.*

Proof: The two calls to Match in stage i are intended to maintain the two claimed properties: that v 's degree degrades geometrically in each round and that the sum of v 's neighbors' degrees degrades geometrically. The proof is by induction on i ; the base case is trivial. For the sake of minimizing notation we use deg_i, Γ_i , etc. to refer to the degree and neighborhood functions just before *each* call to Match in stage i . Consider a vertex $v \in V_i$ at the beginning of stage i . By the inductive hypothesis $\text{deg}_i(v) \leq \delta_{i-1}$ and $\text{deg}_i^{(2)}(v) \leq \nu_{i-1}$, from which it follows that v can have at most $\nu_{i-1}/\tau_i = \frac{\Delta \sqrt{c_1 \ln n}}{2\rho^{i-2}} = \delta_i \cdot (\rho^2/2)$ neighbors *not* in V_i^{lo} . If $\text{deg}_i(v) > \delta_i$ (i.e., $v \in V_i^{\text{hi}}$) then in the

first call to Match, v will be matched with probability⁵ $1 - (1 - 1/\tau_i)^{(1-\rho^2/2)\delta_i} > 1 - e^{-(1-\rho^2/2)c_1 \ln n/2}$. By a union bound all vertices in V_i^{hi} are matched with probability at least $1 - 1/n^{c_1(1-\rho^2/2)/2-1} = 1/\text{poly}(n)$.

We now argue that after the second call to Match, $\deg_{i+1}^{(2)}(v) \leq \nu_i$. Call a node *chosen* in the Match procedure if it has positive indegree in F_1 . A node v will be guaranteed to have positive degree in F_2 if it is chosen *or* if it chooses an edge (v, u) and u has indegree 1 in F_1 , i.e., u has *no choice* but to put (v, u) in F_2 . Once in a path or cycle in F_2 the probability that v is matched is at least $1/2$.

We evaluate the edges chosen by V_i -vertices for F_1 sequentially, beginning with all vertices outside of $\Gamma_i(v)$, then to each vertex in $\Gamma_i(v)$ one at a time, in *descending* order of node ID. (Recall that these were used for tie-breaking in Match.) Let $u \in \Gamma_i(v)$ be the current neighbor under consideration. If at least $\deg_i(u)/2$ neighbors of u are currently unchosen (by vertices already evaluated) then place u in set A , otherwise place u in set B . If u was put in set A and u *does* choose a previously unchosen neighbor (implying that it has positive degree in F_2) then also place u in set A' .

We first analyze the case that $\deg_i(A) \geq \deg_i^{(2)}(v)/2 \geq \nu_i/2$. (If $\deg_i^{(2)}(v) < \nu_i$ there is nothing to prove.) Observe that each vertex u , once in A , is moved to A' with probability at least $1/2$, and if so, contributes $\deg_i(u) \leq \delta_i$ to $\deg_i(A')$.⁶ The probability that after evaluating each $u \in \Gamma_i(v)$, $\deg_i(A')$ is less than half its expectation is:

$$\begin{aligned}
\Pr[\deg_i(A') < \tfrac{1}{2} \mathbb{E}[\deg_i(A')]] &\leq \Pr[\deg_i(A') < \tfrac{1}{4} \deg_i(A)] \\
&\leq \exp\left(-\frac{(\frac{1}{4} \deg_i(A))^2}{2 \sum_{u \in A} (\deg_i(u))^2}\right) && \{\text{Corollary A.4}\} \\
&\leq \exp\left(-\frac{1}{32} \frac{\deg_i(A)^2}{(\deg_i(A)/\delta_i)\delta_i^2}\right) && \{\deg_i(u) \leq \delta_i\} \\
&\leq \exp\left(-\frac{1}{32} \frac{\deg_i(A)}{\delta_i}\right) \\
&\leq \exp\left(-\frac{1}{128} \tau_i\right) && \{\deg_i(A) \geq \nu_i/2 = \delta_i \tau_i/4\} \\
&= \exp\left(-\frac{\Delta}{64\rho^i \sqrt{c_1 \ln n}}\right)
\end{aligned}$$

⁵Note that since $V_i^{\text{lo}} \cap V_i^{\text{hi}} = \emptyset$, F_1 consists of stars and F_2 consists of non-adjacent edges, all of which are added to the matching.

⁶Note that this process fits in the martingale framework of Corollary A.4. Here X_j is the state of the system after evaluating the j th neighbor u of v and Z_j is $\deg_i(u)$ if u joins A' and 0 otherwise, which is a function of X_j . Thus, each Z_j has a range of at most δ_i .

We proceed under the assumption that $\deg_i(A') \geq \frac{1}{4} \deg_i(A) \geq \nu_i/8$. Since each vertex with positive degree in F_2 is matched with probability at least $1/2$, $\mathbb{E}[\deg_{i+1}(A')] \leq \frac{1}{2} \deg_i(A')$. Moreover, whether $v \in A'$ is matched depends only on the bits selected by its neighbors in F_2 , that is, the dependency graph of these events has chromatic number $\chi = 3$. Thus,

$$\begin{aligned}
\Pr[\deg_{i+1}(A') > \tfrac{1}{4} \deg_i(A')] &\leq \exp\left(-\frac{2(\frac{1}{4} \deg_i(A'))^2}{\chi \cdot \sum_{u \in A'} (\deg_i(u))^2}\right) && \{\text{Theorem A.2}\} \\
&\leq \exp\left(-\frac{1}{24} \frac{\deg_i(A')^2}{(\deg_i(A')/\delta_i)\delta_i^2}\right) && \{\chi = 3, \deg_i(u) \leq \delta_i\} \\
&\leq \exp\left(-\frac{1}{24} \frac{\deg_i(A')}{\delta_i}\right) \\
&\leq \exp\left(-\frac{1}{24} \frac{\tau_i}{16}\right) && \{\deg_i(A') \geq \nu_i/8 = \delta_i \tau_i/16\} \\
&\leq \exp\left(-\frac{\Delta}{192\rho^i \sqrt{c_1 \ln n}}\right)
\end{aligned}$$

The second case to consider is if $\deg_i(B) \geq \deg_i^{(2)}(v)/2 \geq \nu_i/2$. Note that $\mathbb{E}[\deg_{i+1}(B)] \leq \frac{3}{4} \deg_i(B)$ since at least half of B -vertices' neighbors are matched with probability at least $1/2$. By another application of Theorem A.2 we have

$$\begin{aligned}
\Pr[\deg_{i+1}(B) > \tfrac{7}{8} \deg_i(B)] &\leq \exp\left(-\frac{2(\frac{1}{8} \deg_i(B))^2}{\chi \cdot \sum_{u \in B} \deg_i(u)^2}\right) && \{\text{Theorem A.2}\} \\
&\leq \exp\left(-\frac{1}{96} \frac{\deg_i(B)^2}{(\deg_i(B)/\delta_i)\delta_i^2}\right) && \{\chi = 3, \deg_i(u) \leq \delta_i\} \\
&\leq \exp\left(-\frac{1}{96} \frac{\tau_i}{4}\right) && \{\deg_i(B) \geq \nu_i/2 = \delta_i \tau_i/4\} \\
&= \exp\left(-\frac{\Delta}{192\rho^i \sqrt{c_1 \ln n}}\right)
\end{aligned}$$

By a union bound, $\deg_{i+1}^{(2)}(v) \leq \frac{7}{8} \deg_i^{(2)}(v)$ for all $v \in V$ with probability $1 - n \exp(-\Omega(\tau_i))$. We set $\rho = \sqrt{8/7} \approx 1.06$, concluding the induction. \square

Lemma 3.1 implies that after $l = O(\ln(\Delta/\ln^{3/2} n))$ stages the maximum degree is at most $\delta_l = (c_1 \ln n/2)\tau_l = O(\ln^2 n)$. Lemma 3.2 implies that by the end of Phase I all surviving connected components have size $\text{poly}(\log n)$.

Lemma 3.2 *At any point in Phase I, if the maximum degree in the graph induced by unmatched vertices is $\tilde{\Delta}$, then for some constant c_4 , after $c_4 \log \tilde{\Delta}$ stages all connected components of unmatched vertices have size $\tilde{\Delta}^4 \log n$, with probability $1 - 1/\text{poly}(n)$.*

Proof: The observations made in Lemma 3.1 imply that in each call to $\text{Match}(V_i, V_i)$, each u loses a constant fraction of its neighbors (either because they are matched or u itself is matched) with constant probability. Moreover, the event that this occurs (a *success* for u) is independent of the success or failure of any u' at distance at least 5 from u .

We use the approach of [6, 29] to show that no components with size $> \tilde{\Delta}^4 \log n$ survive $c_4 \log \tilde{\Delta}$ stages. Consider a subgraph H of G with s vertices. One can easily see that there is some $V_0(H) \subseteq V(H)$ with $|V_0(H)| \geq s/\tilde{\Delta}^4 = t$ such that for all $u, u' \in V_0(H)$, $\text{dist}(u, u') \geq 5$ and $\text{dist}(u, V_0(H) \setminus \{u\}) = 5$.⁷ Such a set $V_0(H)$ corresponds to a tree with size t in the graph $G^5 = (V, \{(u, u') \mid \text{dist}(u, u') = 5\})$, which has maximum degree less than $\tilde{\Delta}^5$. There are fewer than

⁷For example, repeatedly select a vertex u in $V(H)$ at distance 5 from some previously selected vertex, then remove all vertices within distance 4 of u .

4^t distinct trees on t vertices and fewer than $n \cdot \tilde{\Delta}^{5t}$ ways to embed a tree on t vertices in G^5 . For any one vertex the probability that it is not eliminated is at most the probability that it is not successful $O(\log \tilde{\Delta})$ times after $c_4 \log \tilde{\Delta}$ stages, which can be made $\tilde{\Delta}^{-c_5}$ for any c_5 by making c_4 sufficiently large. Since $V_0(H)$ -vertices are at distance at least 5 from each other, these events are independent and the probability that H survives $c_4 \log \tilde{\Delta}$ stages is at most $\tilde{\Delta}^{-c_5 t}$. By a union bound, the probability that any such H survives is at most $n \cdot 4^t \cdot \tilde{\Delta}^{5t - c_5 t} = 1/\text{poly}(n)$ for $t = \log n$. \square

Theorem 3.3 *A maximal matching can be computed in $O(\log \Delta + \log^4 \log n)$ time w.h.p. in an arbitrary distributed network.*

Proof: Lemmas 3.1 and 3.2 imply that after the $c_2 \log \Delta$ stages of Phase I, for a sufficiently large constant c_2 , all connected components of unmatched edges have size $s = O(\log^9 n)$ with probability $1 - 1/\text{poly}(n)$. Phase II is implemented with the deterministic maximal matching algorithm of [10], which runs in $O(\log^4 s) = O(\log^4 \log n)$ time. \square

4 Maximal Independent Set Algorithms

To compute an MIS efficiently we employ the same general strategy used in our MM algorithm. We run a randomized algorithm (a variant of Luby's in this case) for a certain amount of time then argue that the connected components in the graph induced by vertices with degree at least $\Delta/2$ have weak diameter $O(\sqrt{\log n})$ (variant 1), or if $\Delta = \text{poly}(\log n)$, have size $O(\text{poly}(\log n))$ (variant 2). In the first case we use the trivial $O(\text{weak diameter}) = O(\sqrt{\log n})$ MIS algorithm and in the second we use the Panconesi-Srinivasan [28] algorithm, which runs in $2^{O(\sqrt{\log \log n})}$ time. Applying this halving algorithm $\log \Delta$ times reduces the maximum degree to zero. Since $O(\log \Delta \sqrt{\log n})$ is not an improvement over the logarithmic time MIS algorithms for $\Delta > 2^{\sqrt{\log n}}$, we assume in this section that $\Delta \leq 2^{\sqrt{\log n}}$.

Phase I of Halve computes independent sets $I_0 = \emptyset \subseteq I_1 \subseteq \dots \subseteq I_\kappa$ and Phase II computes an MIS on the components of high-degree ($\geq \Delta/2$) vertices in $V \setminus \hat{\Gamma}(I_\kappa)$. In stage i the *active* vertices are $A_i = V \setminus \hat{\Gamma}(I_i)$ and \deg_i and Γ_i are the functions with respect to A_i .

Halve—Phase I:

Initialize $I_0 \leftarrow \emptyset$ and execute stages $0, 1, \dots, \kappa - 1$ where $\kappa = c_6 \sqrt{\log n}$ (variant 1) or $c_6 \log \Delta$ (variant 2).

Stage i :

1. Each $v \in A_i$ selects itself with probability $\frac{1}{\Delta+1}$.
2. $I_{i+1} \leftarrow I_i \cup \{v \mid v \text{ is the only vertex in } \hat{\Gamma}_i(v) \text{ that selects itself}\}$.

Halve—Phase II:

Let $U = \{v \in A_\kappa : \deg_\kappa(v) \geq \Delta/2\}$ be the set of active vertices with degree at least $\Delta/2$. Let \mathcal{C} be the set of connected components of $G(U)$ with weak diameter less than $5\sqrt{\log n}$ (variant 1) or size less than $\Delta^4 \log n$ (variant 2). Deterministically compute an MIS $I(C)$ for each $C \in \mathcal{C}$ and return $I_\kappa \cup \bigcup_{C \in \mathcal{C}} I(C)$.

Lemma 4.1 *Let $S \subseteq A_i$ be such that $\text{dist}(u, u') \geq 5$ and $\deg_i(u) \geq \Delta/2$ for all $u, u' \in S$. The probability that $S \subseteq A_{i+1}$ is less than $p^{|S|}$ where $p = 1 - (1 - e^{-1/2})e^{-1} \approx 0.85$.*

Proof: Whether a vertex becomes inactive in stage i depends only on the choices of vertices within distance 2. The distance-5 constraint ensures that the events $\{u \in A_{i+1}\}_{u \in S}$ are independent, so it suffices to show that $\Pr[u \in A_{i+1}] \leq p$. Let $\{u = v_0, v_1, v_2, \dots, v_d\} = \hat{\Gamma}_i(u)$ be the set containing $u = v_0 \in S$ and its active neighbors v_1, \dots, v_d and let j be the first index such that v_j selects itself in stage i . The probability that j exists is $1 - (1 - \frac{1}{\Delta+1})^{d+1} > 1 - e^{-\frac{\Delta/2+1}{\Delta+1}} > 1 - e^{-1/2}$, and if it exists it will join I_{i+1} if none of its neighbors select themselves. By definition v_0, \dots, v_{j-1}

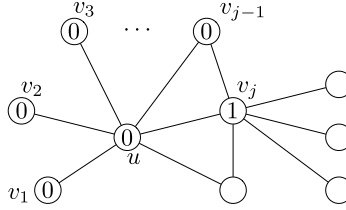


Figure 2: The vertex v_j is the first among $\{u = v_0, \dots, v_d\}$ to select itself. The probability that all neighbors of v_j will not select themselves is: $(1 - 1/(\Delta + 1))^{\Gamma_i(v_j) \setminus \{v_0, \dots, v_{j-1}\}}$.

have *not* selected themselves, so the probability that v_j joins the MIS (and makes u inactive) is $(1 - \frac{1}{\Delta+1})^{|\Gamma_i(v_j) \setminus \{v_0, \dots, v_{j-1}\}|} \geq (1 - \frac{1}{\Delta+1})^\Delta > e^{-1}$. Thus, $\Pr[u \notin A_{i+1}] > (1 - e^{-1/2})e^{-1}$. See Figure 2 for an illustration. \square

In order to prove that U induces components with weak diameter less than $5\sqrt{\log n}$ it suffices to prove that for each $u, u' \in V$ at distance at least $5(\sqrt{\log n} - 1)$, every path from u to u' contains some vertex not in U . To that end we define \mathcal{P} to be the set of *all* paths (not necessarily shortest) between pairs of vertices at distance at least $5(\sqrt{\log n} - 1)$. We first claim that each $P = (u_1, \dots, u_r) \in \mathcal{P}$ contains a $Q(P) = \{q_1, \dots, q_{\sqrt{\log n}}\} \subset V(P)$ such that $\text{dist}(u, u') \geq 5$ for all $u, u' \in Q(P)$. We generate $Q(P)$ one vertex at a time maintaining the invariant that $\text{dist}(q_j, u_r) \geq 5(\sqrt{\log n} - j)$. Define $q_1 = u_1$ and once q_j is known, define $q_{j+1} = u_k$ where k is the maximum index such that $\text{dist}(q_j, u_k) = 5$. It follows that for all $j' < j$, $\text{dist}(q_{j'}, q_{j+1}) > 5$. By the triangle inequality, $\text{dist}(q_j, u_r) \geq 5(\sqrt{\log n} - j)$ implies $\text{dist}(q_{j+1}, u_r) \geq 5(\sqrt{\log n} - (j + 1))$.

Define $\mathcal{Q} = \{Q(P) \mid P \in \mathcal{P}\}$ and \mathcal{W} to be the set of all walks of length exactly $5(\sqrt{\log n} - 1)$. Every $Q(P) \in \mathcal{Q}$ can be mapped injectively to a walk in \mathcal{W} by taking the concatenation of arbitrary shortest paths between successive vertices in $Q(P)$. Thus, $|\mathcal{Q}| \leq |\mathcal{W}| \leq n\Delta^{5(\sqrt{\log n} - 1)}$.

Lemma 4.2 *After $\kappa = c_6\sqrt{\log n}$ (or $\kappa = c_6 \log \Delta$) stages of Phase I, U induces components with weak diameter less than $5\sqrt{\log n}$ (or size less than $\Delta^4 \log n$) w.h.p., for a sufficiently large c_6 .*

Proof: If the weak diameter criterion is violated then there is some $P \in \mathcal{P}$ with $V(P) \subseteq U$. By Lemma 4.1, in each stage i that all $Q(P)$ -vertices have degree at least $\Delta/2$ the probability that none become inactive in stage $i + 1$ is $\exp(-\Omega(|Q(P)|)) = \exp(-\Omega(\sqrt{\log n}))$. Thus, after κ iterations the probability that $Q(P) \subseteq U$ is $\epsilon = \exp(-\Omega(\sqrt{\log n}) \cdot \kappa) = \exp(-\Omega(c_6 \log n))$. By a union bound, the probability that any $Q \in \mathcal{P}$ has $Q(P) \subseteq U$ is $\epsilon |\mathcal{Q}| < \epsilon n\Delta^{5(\sqrt{\log n} - 1)} < \epsilon n^6$ (since $\Delta \leq 2^{\sqrt{\log n}}$), which is $1/\text{poly}(n)$ for sufficiently large c_6 . It follows that no $P \in \mathcal{P}$ has $V(P) \subseteq U$ with probability $1 - 1/\text{poly}(n)$ and that Phase II successfully makes inactive all vertices in U .

The analysis of the second variant of the algorithm follows that of Lemma 3.2. Each subgraph with s vertices contains a vertex set with size $t \geq s/\Delta^4$ that forms a tree in $G^5 = (V, \{(u, u') \mid \text{dist}(u, u') = 5\})$ with size t . There are at most $n4^t\Delta^{5(t-1)}$ trees embedded in G^5 with size t . If each of the t vertices has degree $\Delta/2$ in stage i , by Lemma 4.1 the probability that all t are active in stage $i+1$ is $\exp(-\Omega(t))$ and the probability that all are in U is $\epsilon = \exp(-\Omega(t\kappa)) = \exp(-\Omega(c_6 t \log \Delta))$. By a union bound the probability that a component with size $s = t\Delta^4$ exists in U is less than $\epsilon n 4^t \Delta^{5(t-1)}$, which is $1/\text{poly}(n)$ for $t = \log n$ and sufficiently large c_6 . \square

Theorem 4.3 *An MIS can be computed in $O(\log \Delta \sqrt{\log n})$ time w.h.p. in an arbitrary distributed network, or in $\exp(O(\sqrt{\log \log n}))$ time w.h.p. when $\Delta = \text{poly}(\log n)$.*

Proof: The first bound follows from Lemma 4.2 and the fact that an MIS $I(C)$ for $C \in \mathcal{C}$ can be computed in $O(\sqrt{\log n})$ time in Phase II as follows. The vertex $v \in C$ with minimum ID ascertains $G(C)$ in $O(\sqrt{\log n})$ time, selects an arbitrary MIS, and broadcasts its choice to all vertices in C , also in $O(\sqrt{\log n})$ time. The second bound follows from Lemma 4.2 by applying the Panconesi-Srinivasan algorithm in Phase II, which runs in $\exp(O(\sqrt{\log(\Delta^4 \log n)})) = \exp(O(\sqrt{\log \log n}))$ time for $\Delta = \text{poly}(\log n)$. \square

5 A $(\Delta + 1)$ -Coloring Algorithm

Schneider and Wattenhofer [32] presented a randomized $(\Delta + 1)$ -coloring algorithm running in $O(\log \Delta + \sqrt{\log n})$ time and several faster $O(\Delta)$ -coloring algorithms assuming $\Delta = \Omega(\log n)$. Here we give a faster $(\Delta + 1)$ -coloring algorithm running in $O(\log \Delta + \exp(O(\sqrt{\log \log n})))$ time, which also implies that a graph can be $O(\Delta)$ -colored in $\exp(O(\sqrt{\log \log n}))$ time, for any Δ .⁸

Theorem 5.1 *A $(\Delta + 1)$ -coloring can be computed in $O(\log \Delta + \exp(O(\sqrt{\log \log n})))$ time w.h.p. in an arbitrary distributed network.*

We analyze the following natural randomized coloring algorithm [13]. Let $\Psi = \{1, \dots, \Delta + 1\}$ be the palette. Let $c_i : V \rightarrow \Psi \cup \{\perp\}$ be the partial coloring before the i th stage, where \perp indicates no color, and let $\Gamma_i(v) = \{u \in \Gamma(v) \mid c_i(u) = \perp\}$ be the uncolored neighborhood of v . In the i th stage each colored vertex retains its color and each uncolored vertex v selects a color $c'(v)$ uniformly at random from its available palette $\Psi_i(v) \stackrel{\text{def}}{=} \Psi \setminus \{c(u) \mid u \in \Gamma(v)\}$. It sets $c_{i+1}(v) = c'(v)$ if $c'(v) \notin \{c'(u) \mid u \in \Gamma_i(v)\}$.

From v 's point of view a color $q \in \Psi_i(v)$ is *available* if no neighbor of v selects q . We will call v *happy* if at least $|\Psi_i(v)|/8 \geq (\deg_i(v) + 1)/8$ of its colors are available. The probability that a happy v becomes colored in stage i is therefore at least $1/8$.

Let $\Psi^{-1}(q) \stackrel{\text{def}}{=} \{u \in \Gamma_i(v) \mid q \in \Psi_i(u)\}$ denote the neighbors of v vying for color q and let $w(q) \stackrel{\text{def}}{=} \sum_{u \in \Psi^{-1}(q)} 1/|\Psi_i(u)|$ be the *weight* on color q , i.e., each u distributes $1/|\Psi_i(u)|$ units of weight to each color in its palette. Note that since $\deg_i(u) \geq 1$ for $u \in \Gamma_i(v)$, $|\Psi_i(u)| \geq 2$ and therefore $w(q) \leq |\Psi^{-1}(q)|/2$. We can bound the probability that a particular color q is available as a function of $w(q)$.

⁸If $\Delta > \log n$ use Schneider-Wattenhofer [32]; if $\Delta < \log n$ use our algorithm.

$$\begin{aligned}
\Pr[q \text{ is available to } v] &= \prod_{u \in \Psi^{-1}(q)} \left(1 - \frac{1}{|\Psi_i(u)|}\right) \\
&\geq \left(1 - \frac{w(q)}{|\Psi^{-1}(q)|}\right)^{|\Psi^{-1}(q)|} \quad \{\text{min. when } \Psi^{-1}(q) \text{ contribute equal wt.}\} \\
&= \left(\left(1 - \frac{w(q)}{|\Psi^{-1}(q)|}\right)^{|\Psi^{-1}(q)|/w(q)}\right)^{w(q)} \\
&\geq \left(\frac{1}{4}\right)^{w(q)} \quad \{w(q)/|\Psi^{-1}(q)| \leq 1/2\}
\end{aligned}$$

The expected number of colors available to v is therefore at least

$$\begin{aligned}
\min_w \sum_{q \in \Psi_i(v)} \left(\frac{1}{4}\right)^{w(q)} &\geq \min_w |\Psi_i(v)| \cdot \left(\frac{1}{4}\right)^{\sum_{q'} w(q')/\deg_i(v)} \quad \{\text{minimized when all weights equal}\} \\
&\geq |\Psi_i(v)|/4 \quad \{\sum_{q'} w(q') \leq \deg_i(v)\} \\
&= (\deg_i(v) + 1)/4
\end{aligned}$$

If the number of available colors is at least half its expectation then v is happy. Since the indicator variables for the availability of colors are negatively correlated,⁹ Theorem A.1 bounds the probability that v is *not* happy at $\exp(-2(\deg_i(v)/8)^2/\deg_i(v)) = \exp(-\deg_i(v)/32)$. Let $d^* = 32c_7 \ln n$, for any desired constant c_7 , and let $V_i^* = \{v \in V_i \mid \deg_i(v) \geq d^*\}$. It follows that all of V_i^* is happy with probability $1 - 1/n^{c_7-1}$; call this event H .

Let $\deg_i^*(v) = \{u \in \Gamma_i(v) \mid u \in V_i^*\}$ be the number of neighbors of v with degree greater than d^* . Conditioning on H , we evaluate whether each $u \in (\Gamma_i(v) \cap V_i^*)$ is colored in stage i . The probability that u is colored depends on previously evaluated neighbors of v but is nonetheless always at least $1/8$, due to H . By Corollary A.4 the probability that $\deg_{i+1}^*(v) \geq (\frac{15}{16})\deg_i^*(v)$ is at most $\exp(-(\deg_i^*(v)/16)^2/(2\deg_i^*(v))) = \exp(-\deg_i^*(v)/512)$. In other words, after $r = \log_{16/15} \Delta$ stages, no vertex has $512d^*$ uncolored neighbors each with d^* uncolored neighbors, with probability $1 - 1/n^{c_7-O(1)}$. This does not imply any upper bound on the maximum degree, but we can reduce the remaining coloring problem to two problems with bounded maximum degree. Let $V_{<} \subseteq V_{r+1}$ be the vertices with degree less than d^* . Following the proofs of Theorems 3.3 and 4.3, it is easy to show that after $O(\log d^*) = O(\log \log n)$ more stages of coloring, the connected components induced by uncolored vertices in $V_{<}$ have size $(d^*)^4 \log n = \text{poly}(\log n)$, and can therefore be colored deterministically in $\exp(O(\sqrt{\log \log n}))$ time. The remaining uncolored graph necessarily has maximum degree $512d^*$, and, similarly, will decompose into uncolored components of size $(d^*)^4 \log n$ after $O(\log d^*) = O(\log \log n)$ more stages of coloring. In total the time is $O(\log \Delta + \exp(O(\sqrt{\log \log n})))$.

6 Bounded Arboricity Graphs

Recall that a graph has arboricity λ if its edge set is the union of λ forests. In the proofs of Lemma 6.1 and Theorem 6.2, $\deg_{E'}(u)$ is the number of edges incident to u in $E' \subseteq E$ and $\deg_{V'}(u) = \deg_{G(V' \cup \{u\})}(u)$ is the number of neighbors of u in $V' \subseteq V$.

⁹See Dubhashi and Ranjan [8] for an elementary proof that all balls-and-bins problems of this type give rise to negatively correlated variables. In this case the bins are colors and the balls are the color choices of v 's neighbors.

Lemma 6.1 *Let G be a graph of m edges, n vertices, and arboricity λ .*

1. $m \leq \lambda n$.
2. *The number of vertices with degree at least $t \geq \lambda + 1$ is at most $\lambda n / (t - \lambda)$.*
3. *The number of edges whose endpoints both have degree at least $t \geq \lambda + 1$ is at most $\lambda m / (t - \lambda)$.*

Proof: Part 1 follows from the definition of arboricity. For Parts 2 and 3, let $U = \{v \mid \deg(v) \geq t\}$ be the set of high-degree vertices. Since $|E(U)| \leq \lambda|U|$ (by Part 1) and the sum of degrees of U -vertices is at least $t|U|$, it follows that $|E(U, V \setminus U)| \geq (t - 2\lambda)|U|$. Observe that the number of edges that cross between U and $V \setminus U$ is at least $\sum_{u \in U} (t - \deg_U(u))$. Also, the number of edges $|E(U)|$ induced by U is $\frac{1}{2} \sum_{u \in U} \deg_U(u)$. It follows that the overall number m of edges in the graph satisfies

$$m \geq \sum_{u \in U} (t - \deg_U(u)) + \frac{1}{2} \sum_{u \in U} \deg_U(u) \geq t \cdot |U| - |E(U)| \geq (t - \lambda) \cdot |U|.$$

Hence $\lambda n \geq m \geq (t - \lambda) \cdot |U|$, and so $|U| \leq \lambda n / (t - \lambda)$, proving part 2. Part 3 follows since the number of such edges is at most $\lambda|U| \leq \lambda m / (t - \lambda)$. \square

Theorem 6.2 *Let G be a graph of arboricity λ and $t \geq \max\{\lambda^8, (4(c+1)\ln n)^7\}$ be a parameter. In $O(\log_t n)$ time we can find a matching $M \subseteq E(G)$ (or an independent set $I \subseteq V(G)$) such that with probability at least $1 - 1/n^c$, the maximum degree in the induced graph $G(V \setminus V(M))$ (or $G(V \setminus \hat{I}(I))$) is at most $t\lambda$.*

Proof: In $O(\log_t n)$ rounds we commit edges to M (or vertices to I) and remove all incident edges (or incident vertices). Let G be the graph still under consideration before some round and let $\mathcal{H} = \{v \in V \mid \deg_G(v) \geq t\lambda\}$ be the remaining high-degree vertices. Our goal is to reduce the size of \mathcal{H} by a roughly $t^{1/7}$ factor. Let $\mathcal{J} = \{v \in \mathcal{H} \mid \deg_{\mathcal{H}}(v) \geq t\lambda/2\}$. It follows that any vertex $v \in \mathcal{H}' = \mathcal{H} \setminus \mathcal{J}$ has $\deg_{V \setminus \mathcal{H}}(v) \geq t\lambda/2$. Let $\tilde{E} \subset E(\mathcal{H}', V \setminus \mathcal{H})$ be any set of edges such that for $v \in \mathcal{H}'$, $\deg_{\tilde{E}}(v) = t\lambda/2$ (that is, discard all but $t\lambda/2$ edges arbitrarily) and let $\mathcal{S} = \{u \mid v \in \mathcal{H}' \text{ and } (v, u) \in \tilde{E}\}$ be the neighborhood of \mathcal{H}' with respect to \tilde{E} . Note that $|\mathcal{S}| \leq t\lambda|\mathcal{H}'|/2$. See Figure 3. We define bad \mathcal{S} -vertices, bad \tilde{E} -edges, and bad \mathcal{H}' -vertices as follows, where $\beta = t^{1/7}$.

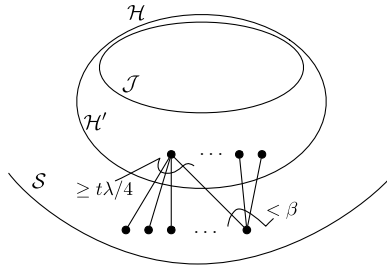


Figure 3: Good \mathcal{S} -vertices have fewer than β neighbors in \mathcal{H}' and fewer than β^2 neighbors in \mathcal{S} . Good \mathcal{H}' -vertices have at least $t\lambda/4$ good neighbors in \mathcal{S} .

Let $B_{\mathcal{S}} = \{u \in \mathcal{S} \mid \deg_{\tilde{E}}(u) \geq \beta \text{ or } \deg_{\mathcal{S}}(u) \geq \beta^2\}$, let $B_{\tilde{E}} = \{(u, v) \in \tilde{E} \mid u \in B_{\mathcal{S}}\}$, and let $B_{\mathcal{H}'} = \{v \in \mathcal{H}' \mid \deg_{\tilde{E} \setminus B_{\tilde{E}}}(v) < t\lambda/4\}$.

By Lemma 6.1(3) the number of bad $(u, v) \in B_{\tilde{E}}$ due to $\deg_{\tilde{E}}(u) \geq \beta$ is at most $\lambda|\tilde{E}|/(\beta - \lambda) \leq \lambda(t\lambda|\mathcal{H}'|/2)/(\beta - \lambda)$. By Lemma 6.1(2) the number of *additional* bad $(u, v) \in B_{\tilde{E}}$ due to $\deg_{\mathcal{S}}(u) \geq \beta^2$ is at most $(\beta - 1)\lambda|\mathcal{S}|/(\beta^2 - \lambda) \leq (\beta - 1)\lambda(t\lambda|\mathcal{H}'|/2)/(\beta^2 - \lambda)$ since there are at most $\lambda|\mathcal{S}|/(\beta^2 - \lambda)$ such u and each contributes fewer than β edges in \tilde{E} . In total $|B_{\tilde{E}}| < \lambda^2 t |\mathcal{H}'|/\beta$. Note that a bad $v \in \mathcal{H}'$ must be incident to more than $t\lambda/4$ edges in $B_{\tilde{E}}$ since $\deg_{\tilde{E}}(v) = t\lambda/2$. Hence $|B_{\mathcal{H}'}| < |B_{\tilde{E}}|/(t\lambda/4) < 4\lambda|\mathcal{H}'|/\beta = 4\lambda|\mathcal{H}'|/t^{1/7}$.

Our goal now is to select some vertices for the MIS (or edges for the MM) that eliminate all good \mathcal{H}' vertices, with high probability. Each vertex $u \in \mathcal{S} \setminus B_{\mathcal{S}}$ selects a random number and joins the MIS if it holds a local maximum. The probability that u joins the MIS is $1/(\deg_{\mathcal{S} \setminus B_{\mathcal{S}}}(u) + 1) \geq 1/\beta^2$ and this event is clearly independent of all $u' \in \mathcal{S} \setminus B_{\mathcal{S}}$ at distance (in $\mathcal{S} \setminus B_{\mathcal{S}}$) at least 3. Note that since the maximum degree in the graph induced by $\mathcal{S} \setminus B_{\mathcal{S}}$ is less than β^2 , the neighborhood of any good $v \in \mathcal{H}' \setminus B_{\mathcal{H}'}$ contains a subset of at least $(t\lambda/4)/\beta^4$ vertices, each pair of which is at distance at least 3 with respect to $\mathcal{S} \setminus B_{\mathcal{S}}$. (Such a set could be selected greedily.) Thus, the probability that no neighbor of $v \in \mathcal{H}' \setminus B_{\mathcal{H}'}$ joins the MIS is at most $(1 - 1/\beta^2)^{t\lambda/(4\beta^4)} < e^{-t\lambda/(4\beta^6)} = e^{-t^{1/7}\lambda/4} < 1/n^{c+1}$. Thus, with high probability all good $\mathcal{H}' \setminus B_{\mathcal{H}'}$ vertices are eliminated. Any remaining high degree vertices must be in \mathcal{J} or $B_{\mathcal{H}'}$, which, by Lemma 6.1(2) number at most $\lambda|\mathcal{H}|/(t/2 - \lambda) + 4\lambda|\mathcal{H}'|/t^{1/7} < 5\lambda|\mathcal{H}|/t^{1/7}$. The number of high-degree vertices is reduced by a $t^{\Omega(1)}$ factor since $t > \max\{\lambda^8, (4(c+1)\ln n)^7\}$, so after $O(\log_t n)$ rounds all high-degree vertices have been eliminated, with probability at least $1 - 1/n^c$.

In the case of MM we want to select a matching that matches all \mathcal{H}' vertices. Each $u \in \mathcal{S} \setminus B_{\mathcal{S}}$ chooses an edge $(u, v) \in \tilde{E} \setminus B_{\tilde{E}}$ uniformly at random and proposes to v that (u, v) be included in the matching. Any $v \in \mathcal{H}' \setminus B_{\mathcal{H}'}$ receiving a proposal accepts one arbitrarily and becomes matched. A good $v \in \mathcal{H}' \setminus B_{\mathcal{H}'}$ has at least $\deg_{\tilde{E} \setminus B_{\tilde{E}}}(v) \geq t\lambda/4$ neighbors $u \in \mathcal{S} \setminus B_{\mathcal{S}}$ with $\deg_{\tilde{E} \setminus B_{\tilde{E}}}(u) < \beta$, so the probability that v receives no proposal (and remains unmatched) is less than $(1 - 1/\beta)^{t\lambda/4} < e^{-t^{6/7}\lambda/4} < o(1/n^{c+1})$. As in the case of MIS, the number of high-degree vertices is reduced by a $t^{\Omega(1)}$ factor in $O(1)$ time. (For the MM problem our proof could be simplified somewhat since edges inside \mathcal{S} play no part in the algorithm and need not be classified as good or bad.) \square

Theorem 6.3 *Given a graph of arboricity λ , an MM can be computed in time on the order of:*

$$\min \left\{ \log \lambda + \sqrt{\log n}, \quad \log \Delta + \lambda + \log \log n \right\}$$

for all λ , and in $O\left(\log \Delta + \frac{\log \log n}{\log \log \log n}\right)$ time for $\lambda < \log^{1-\Omega(1)} \log n$.

Proof: The second and third bounds follow from Theorem 3.3 by substituting for [10] the deterministic MM algorithms of Barenboim and Elkin [3] for small arboricity graphs. Their algorithms run in $O(\frac{\log s}{\log \log s})$ time on graphs with size s and arboricity $\lambda < \log^{1-\Omega(1)} s$ and in time $O(\lambda + \log s)$ in general. In the context of our algorithm, $s \leq \log^9 n$. The first MM bound is a consequence of Theorem 6.2 and Theorem 3.3. We reduce the maximum degree to $\Delta = \lambda t = \lambda \cdot \max\{\lambda^8, 2^{\sqrt{\log n}}\}$ in $O(\log_t n) = O(\sqrt{\log n})$ time and find an MM of the resulting graph in $O(\log \Delta + \log^4 \log n) = O(\log \lambda + \sqrt{\log n})$ time. \square

Theorem 6.4 *Given a graph of arboricity λ , an MIS can be computed in time on the order of:*

$$\min \left\{ \log \lambda \sqrt{\log n} + \log^{3/4} n, \quad \log^2 \Delta + \lambda \log \Delta + \lambda^\epsilon \log \log n, \quad \log^2 \Delta + \lambda^{1+\epsilon} \log \Delta + \log \lambda \log \log n \right\}$$

time for all λ and $\epsilon > 0$, and in $O\left(\log \Delta \left(\log \Delta + \frac{\log \log n}{\log \log \log n}\right)\right)$ time for $\lambda < \log^{1/2-\Omega(1)} \log n$.

Proof: The first bound is a consequence of Theorem 6.2 and Theorem 4.3 since we can reduce the maximum degree to $\Delta = 2^{\log^{1/4} n} \text{poly}(\lambda)$ in $O(\log^{3/4} n)$ time and then find an MIS of the resulting graph in $O(\log \Delta \sqrt{\log n}) = O(\log \lambda \sqrt{\log n} + \log^{3/4} n)$ time.

To obtain the second and third bounds we use the fact that a c -coloring can be converted to an MIS in $O(c)$ time. In Phase II of the Halve algorithm from Section 4 the remaining components have size at most $s = \Delta^4 \log n$. We run one of the deterministic $f(\lambda)$ -coloring algorithms of Barenboim and Elkin [5] on each component. We use a different palette for each call to Halve, so these coloring routines can be pipelined, i.e., Phase I of the next iteration of Halve can start immediately. After $\log \Delta$ calls to Halve every vertex is either in the MIS under construction, adjacent to an MIS vertex (and therefore inactive), or colored one of $f(\lambda) \log \Delta$ colors. Thus, an MIS can be found in an additional $f(\lambda) \log \Delta$ time. For $f(\lambda) = O(\lambda)$ and $f(\lambda) = \lambda^{1+\epsilon}$, respectively, the [5] algorithms run in $O(\lambda^\epsilon \log s)$ time and $O(\log \lambda \log s)$ time.

The fourth bound reflects the use of the Barenboim-Elkin [3] MIS algorithm in Phase II of Halve, which runs in $O(\log s / \log \log s)$ time if $\lambda < \log^{1/2-\Omega(1)} s$. \square

Theorem 6.5 *Given a graph of arboricity λ and any fixed $\epsilon > 0$, a $(\Delta + \lambda^{1+\epsilon})$ -coloring can be computed in $O(\log \Delta + \log \lambda \log \log n)$ time and a $(\Delta + O(\lambda))$ -coloring can be computed in $O(\log \Delta + \lambda^\epsilon \log \log n)$ time. Consequently, a $(\Delta + 1)$ -coloring can be computed in time*

$$O(\log \Delta + \min\{\lambda^{1+\epsilon} + \log \lambda \log \log n, \lambda + \lambda^\epsilon \log \log n\})$$

Proof: Using a palette of $\Delta + 1$ colors, we reduce the size of uncolored components to $s = \text{poly}(\log n)$ in $O(\log \Delta)$ time, as in Section 5, then run one of the Barenboim-Elkin [5] deterministic $f(\lambda)$ -coloring algorithms on each component, which takes $O(\log \lambda \log s)$ time for $f(\lambda) = \lambda^{1+\epsilon}$ and $O(\lambda^\epsilon \log s)$ time for $f(\lambda) = O(\lambda)$. In an additional $O(f(\lambda))$ time we can convert this coloring to a $(\Delta + 1)$ -coloring. \square

7 MIS in Trees and High Girth Graphs

The MIS algorithm of Métivier et al. [24] works as follows. In each round each remaining vertex chooses a random number¹⁰ and includes itself in the MIS if it holds a local maximum, thereby *eliminating* it and its neighborhood from further consideration. Observe that the probability that v joins the MIS in a round is $1/(\deg(v) + 1)$, irrespective of the degrees of its neighbors.

Lemma 7.1 both generalizes a lemma due to Lenzen and Wattenhofer [19] (from trees to high-girth graphs) and corrects a small flaw in their claim and proof. Lemma 7.1 is weaker than [19, Lemma 4.8] by a $\log \log n$ factor, which provides an $O(\sqrt{\log n} \log \log n)$ bound on the running time of [19], rather than the claimed $O(\sqrt{\log n \log \log n})$. Theorem 7.3 improves this to $O(\min\{\sqrt{\log n \log \log n}, \log \Delta \log \log \Delta + \frac{\log \log n}{\log \log \log n}\})$.

Lemma 7.1 *Let $G = (V, E)$ be a graph with girth greater than 6 and maximum degree Δ and let G_i be the remaining graph after $i \cdot c \log \log n$ iterations of the Métivier et al. MIS algorithm. Then for all $v \in V$ and integer $\delta \geq 1$, $|\{w \in \Gamma_i(v) \mid \deg_i(w) > \Delta/2^{i-\delta}\}| < \max\{\Delta/2^{i+\delta}, c' \log n\}$ with high probability, where c and c' are constants, and Γ_i and \deg_i are short for Γ_{G_i} and \deg_{G_i} .*

¹⁰The analysis supposes that vertices choose reals from any common distribution. This can be simulated, for our purposes, by selecting an integer from $[1, \dots, n^b]$ uniformly at random, for a sufficiently large constant b .

Proof: The lemma holds trivially when $i \leq 1$. Assume, inductively, that the lemma holds for $i - 1$ and $\delta \geq 1$. Consider an arbitrary neighbor w of v with $\deg_{i-1}(w) > \Delta/2^{i-\delta}$. According to the hypothesis, $|\{w' \in \Gamma_{i-1}(w) \mid \deg_{i-1}(w') > \Delta/2^{(i-1)-\delta}\}| < \Delta/2^{i-1+\delta} \leq \Delta/2^{i-\delta+1}$. Thus, w has at least $\Delta/2^{i-\delta+1}$ neighbors (beside v) with degree at most $\Delta/2^{(i-1)-\delta}$. Due to the lower bound on girth, the neighborhoods of w 's neighbors intersect only at w . The probability that w is eliminated is at least the probability that some neighbor w' with $\deg_{i-1}(w') \leq \Delta/2^{i-1-\delta}$ joins the MIS. Consider the random numbers $r(\cdot)$ selected by $\hat{\Gamma}_{i-1}(w)$ in one iteration of the MIS algorithm, and let $N = \{w' \in \Gamma_{i-1}(w) \mid \deg_{i-1}(w') \leq \Delta/2^{i-1-\delta} \text{ and } r(w') > r(w)\}$. The events that vertices in N join the MIS are independent as they no longer depend on w . If we condition on $|N| \geq \Delta/2^{i-\delta+2}$, which occurs with probability at least $1/2$, the probability that w is eliminated is at least the probability that some N -vertex joins the MIS, which is at least

$$\begin{aligned} & \frac{1}{2} \left(1 - (1 - 2^{i-1-\delta}/\Delta)^{\Delta/2^{i-\delta+2}} \right) \\ & > \frac{1}{2}(1 - e^{-1/8}) > 1/20 \end{aligned}$$

Now consider v . If $\delta > 1$ then, by the inductive hypothesis, the number of neighbors with degree greater than $\Delta/2^{i-\delta} = \Delta/2^{(i-1)-(\delta-1)}$ is less than $\max\{\Delta/2^{i+\delta-2}, c' \log n\}$ and we need only reduce this number by a factor of 4 (or less, if $\Delta/2^{i+\delta}$ is already less than $c' \log n$.) Since such a neighbor is eliminated with constant probability in each iteration of the algorithm (independently, due to the girth assumption), the probability that v still has more than $\max\{\Delta/2^{i+\delta}, c' \log n\}$ such neighbors after $O(1)$ iterations is, by a Chernoff bound, $\exp(-\Omega(\max\{\frac{\Delta}{2^{i+\delta}}, c' \log n\})) < 1/\text{poly}(n)$.

However, if $\delta = 1$ then the inductive assumption does not apply. We distinguish two cases: either $|\{w \in \Gamma_{i-1}(v) \mid \deg_{i-1}(w) > \Delta/2^{i-\delta}\}|$ is greater than $T = c'' \ln n \cdot \Delta/2^{i+\delta}$ (for a suitably large constant c'') or not. In the first case, we argue that with high probability, v is eliminated in *one* iteration of the Métivier et al. algorithm, slightly modified as follows. Each vertex x whose degree is in the range $[\Delta/2^{i-\delta}, \Delta/2^{i-\delta-1}] = [\Delta/2^{i-1}, \Delta/2^{i-2}]$ independently chooses a random number $r(x)$; other x implicitly choose $r(x) = -\infty$. If x holds a local maximum it joins the MIS. By the inductive hypothesis v has less than $\Delta/2^{(i-1)+\delta} = \Delta/2^i$ neighbors with degree greater than $\Delta/2^{(i-1)-\delta} = \Delta/2^{i-2}$, so at least $T - \Delta/2^i = (c'' \ln n - 2)\Delta/2^{i+1}$ of v 's neighbors will be choosing random numbers. The probability that *none* are local maxima (and therefore v survives) is at most $(1 - 2^{i-2}/\Delta)^{(c'' \ln n - 2)\Delta/2^{i+1}} = 1/\text{poly}(n)$.

In the second case v has fewer than T neighbors with degree at least $\Delta/2^{i-\delta}$. On the other hand, we can assume it has more than $c' \log n$ such neighbors, otherwise the claim is trivially satisfied. Each such neighbor is eliminated with constant probability in each iteration of the algorithm, and these events are independent due to the lower bound on girth. By a Chernoff bound, after $O(\log \log n)$ iterations v will have at most $\Delta/2^{i+\delta}$ such neighbors with probability $1 - \exp(-\Omega(\Delta/2^{i+\delta})) < 1/\text{poly}(n)$. \square

Theorem 7.2 *An MIS can be computed in $O(\log \Delta \log \log n + \exp(O(\sqrt{\log \log n})))$ time on graphs with girth greater than 6.*

Proof: Lemma 7.1 implies that in $O(\log \Delta \log \log n)$ iterations the graph is reduced to a situation where no vertex has $c' \log n$ neighbors with degree at least $c' \log n$. Let L be the set of vertices with degree at most $c' \log n$ and H be the remaining vertices. The maximum degree in the graph induced by H is $\Delta' = c' \log n$, so, by Theorem 4.3, we can find an MIS for H in $\exp(O(\sqrt{\log \log n}))$ time, then find an MIS for what remains of L , also in $\exp(O(\sqrt{\log \log n}))$ time. \square

Theorem 7.3 *An MIS of an unoriented tree can be computed in time on the order of*

$$\min \left\{ \sqrt{\log n \log \log n}, \quad \log \Delta \log \log \Delta + \frac{\log \log n}{\log \log \log n}, \quad \Delta + \log^* n \right\},$$

with high probability.

Proof: If the graph has maximum degree greater than $2\sqrt{\log n / \log \log n}$, reduce it to $2\sqrt{\log n / \log \log n}$ in $\sqrt{\log n \log \log n}$ time, using Theorem 6.2.

In the modified Métivier et al. algorithm from Lemma 7.1, all the critical events held with probability $1 - \exp(-\Omega(c' \log n)) = 1 - 1/\text{poly}(n)$. We run the Lemma 7.1 algorithm, substituting $c' \log \Delta$ for $c' \log n$ and $i \cdot c \log \log \Delta$ for $i \cdot c \log \log n$. In order to guarantee that after $i \cdot c \log \log \Delta$ iterations, no vertex has more than $\max\{\Delta/2^{i+\delta}, c' \log \Delta\}$ neighbors with degree greater than $\Delta/2^{i-\delta}$, for any $\delta \geq 1$, we mark as *bad* any vertex that violates this condition and temporarily remove bad vertices from consideration. More precisely, redefine $\Gamma_i(v)$ to be the neighborhood of v after $i \cdot c \log \log \Delta$ iterations, excluding bad vertices and those covered by the MIS under construction. Since bad vertices are removed from consideration, we guarantee that

$$|\{w \in \Gamma_i(v) \mid \deg_i(w) > \Delta/2^{i-\delta}\}| < \max\{\Delta/2^{i+\delta}, c' \log \Delta\} \quad (1)$$

holds with probability 1, for any $\delta \geq 1$. Given that this property is satisfied at v after $i \cdot c \log \log \Delta$ iterations, the probability that it is also satisfied at v after $(i+1) \cdot c \log \log \Delta$ iterations (and is therefore not marked bad) is at least $1 - \exp(-\Omega(c' \log \Delta)) = 1 - 1/\text{poly}(\Delta)$ for sufficiently large constants c and c' . (The analysis is identical to that in Lemma 7.1.) Thus, the probability that a vertex is marked bad after $O(\log \Delta \log \log \Delta)$ iterations can be made at most $1/(8\Delta)$.

According to Eqn. (1), after $O(\log \Delta \log \log \Delta)$ iterations no vertex has $\Delta' = c' \log \Delta$ neighbors with degree greater than Δ' . As in the proof of Theorem 7.2, let H and L be the set of (non-bad) vertices after $O(\log \Delta \log \log \Delta)$ iterations with degree at least Δ' and less than Δ' , respectively. Note that no vertex has Δ' neighbors in H . First compute an MIS on the graph induced by H , in $O(\Delta' + \log^* n) = O(\log \Delta + \log^* n)$ time, then compute an MIS on what remains of L , also in $O(\Delta' + \log^* n)$ time. At this point all non-bad vertices have been removed. Let B be the set of remaining bad vertices. We claim that with high probability B induces a collection of trees with size $O(\log n)$. With respect to the original tree, there are fewer than $n \cdot (4\Delta)^t$ distinct subtrees with size t . (There are fewer than 4^t distinct rooted trees with size t , and at most $n\Delta^t$ ways to embed such a subtree in the original tree.) The probability that any given subtree of size t is contained in B is $1/(8\Delta)^t$. Thus, by a union bound, the probability that any subtree with size at least t is contained in B is at most $n(4\Delta)^t/(8\Delta)^t = 1/\text{poly}(n)$ for $t = \Theta(\log n)$. The Barenboim-Elkin algorithm [3] computes an MIS of a collection of such trees, deterministically, in $O(\log t / \log \log t) = O(\log \log n / \log \log \log n)$ time. \square

References

- [1] N. Alon, L. Babai, and A. Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7(4):567–583, 1986.
- [2] B. Awerbuch, A. V. Goldberg, M. Luby, and S. Plotkin. Network decomposition and locality in distributed computation. In *Proc. of the 30th Symposium on Foundations of Computer Science*, pages 364–369, 1989.

- [3] L. Barenboim, and M. Elkin. Sublogarithmic distributed MIS algorithm for sparse graphs using Nash-Williams decomposition. In *Proc. of the 27th ACM Symp. on Principles of Distributed Computing*, pages 25–34, 2008.
- [4] L. Barenboim, and M. Elkin. Distributed $(\Delta + 1)$ -coloring in linear (in Δ) time. In *Proc. of the 41st ACM Symp. on Theory of Computing*, pages 111–120, 2009. Full version available at <http://arXiv.org/abs/0812.1379v2>.
- [5] L. Barenboim, and M. Elkin. Deterministic distributed vertex coloring in polylogarithmic time. In *Proc. of the 29th ACM Symp. on Principles of Distributed Computing*, pages 410–419, 2010.
- [6] J. Beck. An Algorithmic Approach to the Lovász Local Lemma. I. *Random Struct. Algorithms* 2(4):343–366, 1991.
- [7] D. P. Dubhashi and A. Panconesi. Concentration of Measure for the Analysis of Randomized Algorithms. Cambridge University Press, 2009.
- [8] D. P. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. *Random Struct. Algorithms* 13(2):99–124, 1998.
- [9] B. Gfeller, and E. Vicari. A randomized distributed algorithm for the maximal independent set problem in growth-bounded graphs. In *Proc. of the 26th ACM Symp. on Principles of Distributed Computing*, pages 53–60, 2007.
- [10] M. Hanckowiak, M. Karonski, and A. Panconesi. On the Distributed Complexity of Computing Maximal Matchings. *SIAM J. Discrete Math.* 15(1):41–57, 2001.
- [11] A. Israeli, and A. Itai. A fast and simple randomized parallel algorithm for maximal matching. *Information Processing Letters* 22(2):77–80, 1986.
- [12] S. Janson. Large deviations for sums of partly dependent random variables. *Random Struct. Algorithms* 24(3):234–248, 2004.
- [13] O. Johansson. Simple Distributed $\Delta + 1$ -coloring of Graphs. *Inf. Process. Lett.* 70(5):229–232, 1999.
- [14] K. Kothapalli, and S. Pemmaraju. Distributed graph coloring in a few rounds. In *Proc. of the 30th ACM Symp. on Principles of Distributed Computing*, pages 31–40, 2011.
- [15] K. Kothapalli, C. Scheideler, M. Onus, and C. Schindelhauer. Distributed coloring in $O(\sqrt{\log n})$ bit rounds. In *Proc. of the 20th International Parallel and Distributed Processing Symposium*, 2006.
- [16] F. Kuhn. Weak graph colorings: distributed algorithms and applications. In *Proc. of the 21st ACM Symposium on Parallel Algorithms and Architectures*, pages 138–144, 2009.
- [17] F. Kuhn, T. Moscibroda, and R. Wattenhofer. What cannot be computed locally! In *Proc. of the 23rd ACM Symp. on Principles of Distributed Computing*, pages 300–309, 2004.

- [18] F. Kuhn, T. Moscibroda, and R. Wattenhofer. Local Computation: Lower and Upper Bounds. <http://arXiv.org/abs/1011.5470>, 2010.
- [19] C. Lenzen and R. Wattenhofer. MIS on trees. In *Proc. of the 30th ACM Symp. on Principles of Distributed Computing*, pages 41-48, 2011.
- [20] N. Linial. Distributive graph algorithms: Global solutions from local data In *Proc. of the 28th Annual Symp. on Foundation of Computer Science*, pages 331–335, 1987.
- [21] N. Linial. Locality in distributed graph algorithms. *SIAM Journal on Computing*, 21(1):193–201, 1992.
- [22] M. Luby. A simple parallel algorithm for the maximal independent set problem. *SIAM Journal on Computing*, 15:1036-1053, 1986.
- [23] P. Long and R. Servedio. Martingale Boosting. In *Proc. of the 18th Annual Conference on Computational Learning Theory*, pages 79-94, 2005. Full version available at http://www.cs.columbia.edu/~rocco/Public/mb_fullversion.pdf
- [24] Y. Métivier, J. M. Robson, N. Saheb-Djahromi, and A. Zemmari. An optimal bit complexity randomized distributed MIS algorithm *Distributed Computing*, 23(5-6):331–340, 2011.
- [25] C. Nash-Williams. Decompositions of finite graphs into forests. *J. London Math*, 39:12, 1964.
- [26] D. Peleg. *Distributed Computing: A Locality-Sensitive Approach*. SIAM, 2000.
- [27] A. Panconesi, and R. Rizzi. Some simple distributed algorithms for sparse networks. *Distributed computing*, 14(2):97–100, 2001.
- [28] A. Panconesi, and A. Srinivasan. On the complexity of distributed network decomposition. *Journal of Algorithms*, 20(2):581-592, 1995.
- [29] R. Rubinfeld, G. Tamir, S. Vardi, and Ning Xie. Fast Local Computation Algorithms. In *Proceedings 2nd Symposium on Innovations in Computer Science (ICS)*, pages 223–238, 2011.
- [30] M. Szegedy and S. Vishwanathan. Locality based graph coloring. In *Proceedings of the Twenty-Fifth Annual ACM Symposium on Theory of Computing*, pages 201-207, 1993.
- [31] J. Schneider, and R. Wattenhofer. A log-star distributed Maximal Independent Set algorithm for Growth Bounded Graphs. In *Proc. of the 27th ACM Symp. on Principles of Distributed Computing*, pages 35–44, 2008.
- [32] J. Schneider, and R. Wattenhofer. A new technique for distributed symmetry breaking. In *Proc. of the 29th ACM Symp. on Principles of Distributed Computing*, pages 257-266, 2010.
- [33] M. Wattenhofer and R. Wattenhofer. Distributed Weighted Matching. In *Proceedings 18th International Conference on Distributed Computing (DISC)*, pages 335–348, 2004.

Appendix

A Concentration Inequalities

See Dubhashi and Panconesi [7] for proofs of Theorems A.1–A.3 and related concentration bounds. In our applications of these inequalities we often simplify the sum $\sum_i (a'_i - a_i)^2$ as follows. If $\sum_i (a'_i - a_i) \leq T$ and $\max_i (a'_i - a_i) \leq t$ then $\sum_i (a'_i - a_i)^2 \leq (T/t)t^2 = tT$.

Theorem A.1 (Chernoff with negative correlation) *Let $X = X_1 + \dots + X_n$ be the sum of n random variables, where the $\{X_i\}$ are independent or negatively correlated. Then for any $t > 0$:*

$$\Pr[X \geq \mathbb{E}[X] + t], \Pr[X \leq \mathbb{E}[X] - t] \leq \exp\left(-\frac{2t^2}{\sum_i (a'_i - a_i)^2}\right)$$

where $a_i \leq X_i \leq a'_i$.

Theorem A.2 (Janson) *For $X = X_1 + \dots + X_n$ the sum of n random variables and $t > 0$:*

$$\Pr[X \geq \mathbb{E}[X] + t], \Pr[X \leq \mathbb{E}[X] - t] \leq \exp\left(-\frac{2t^2}{\chi \cdot \sum_i (a'_i - a_i)^2}\right)$$

where $a_i \leq X_i \leq a'_i$ and χ is the fractional chromatic number of the dependency graph $G_X = (\{1, \dots, n\}, \{(i, j) \mid X_i \text{ and } X_j \text{ are not independent}\})$.

Theorem A.3 (Azuma-Hoeffding) *A sequence Y_0, \dots, Y_n is a martingale with respect to X_0, \dots, X_n if Y_i is a function of X_0, \dots, X_i and $\mathbb{E}[Y_i \mid X_0, \dots, X_{i-1}] = Y_{i-1}$. For such a martingale with bounded differences $a_i \leq Y_i - Y_{i-1} \leq a'_i$,*

$$\Pr[Y_n > Y_0 + t], \Pr[Y_n < Y_0 - t] \leq \exp\left(-\frac{t^2}{2 \sum_i (a'_i - a_i)^2}\right)$$

Corollary A.4 *Let $Z = Z_1 + \dots + Z_n$ be the sum of n random variables and X_0, \dots, X_n be a sequence, where Z_i is a function of X_0, \dots, X_i , $\mu_i = \mathbb{E}[Z_i \mid X_0, \dots, X_{i-1}]$, $\mu = \sum_i \mu_i$, and $a_i \leq Z_i \leq a'_i$. Then*

$$\Pr[Z > \mu + t], \Pr[Z < \mu - t] \leq \exp\left(-\frac{t^2}{2 \sum_i (a'_i - a_i)^2}\right)$$

Proof: Define the martingale Y_0, \dots, Y_n w.r.t. X_0, \dots, X_n by $Y_0 = 0$ and $Y_i = Y_{i-1} + Z_i - \mu_i$, then apply Theorem A.3. Note $Y_n - Y_0 = Z - \mu$ and the range of $Y_i - Y_{i-1}$ still has size $a'_i - a_i$. \square